

AS TÉCNICAS ENVOLVIDAS EM ENGENHARIA GENÉTICA: SUAS PERSPECTIVAS E DIFICULDADES.

Luisa Lina Villa *
& Ricardo Renzo Brentani **

RESUMO: A essência da Engenharia Genética está na junção de fragmentos de DNA ou genes de diferentes organismos com o DNA de vetores plasmidiais ou virais, a fim de serem introduzidos e propagados em bactérias ou outras células hospedeiras. Com o objetivo de analisar a estrutura gênica, os processos envolvidos em sua regulação e expressão ou, ainda, para produzir proteínas específicas de interesse médico ou comercial, uma série de técnicas vem sendo descritas. Este artigo resume a maioria dos métodos empregados, conhecidos em conjunto como a *tecnologia do DNA recombinante*.

UNITERMOS: Engenharia Genética, Clonagem, clones, DNA, RNA, genes, plasmídeos, vírus, enzimas de restrição, hibridização, seqüência de DNA, vetores de expressão.

Sabemos que o constituinte celular mais representativo do fenômeno vital são as proteínas. Estas resultam da união, por meio de ligação peptídica, de aminoácidos, dos quais 22 são empregados pela natureza para este processo.

Como bem evidenciam as hemoglobinopatias, a ordem com que estes 22 aminoácidos são polimerizados é absolutamente crítica, para que o polímero resultante tenha significado biológico. Tal rigor presume a necessidade de um código, o qual deve apresentar suas propriedades fundamentais: deve ser estável, para ser acessível à maquinaria celular sempre que houver necessidade de sintetizar novas moléculas de proteína; deve ser transmissível hereditariamente, para que a prole resultante da divisão de um tipo celular qualquer continue competente para a síntese daquelas proteínas cujo conjunto constitui o fenótipo característico daquele tipo celular.

Sabemos igualmente que tal código de fato existe e que sua natureza química é o DNA. Como o DNA se compõe de quatro bases nitrogenadas, A, T, C e G, segue-se que grupos de três nucleotídeos (base-desoxirribose-fosfato) correspondem aos aminoácidos. Como o arranjo de quatro bases em grupos de três com repetição admitem 64 possibilidades, segue-se que o código é degenerado, ocorrendo sinonímia, isto é, mais de um códon (conjunto de três nucleotídeos) correspondendo ao mesmo aminoácido.

A determinação da massa de DNA em diferentes organismos mostra uma proporcionalidade entre esta massa e o estágio evolutivo. O ganho de DNA entre bactérias e mamíferos é de três ordens de magnitude, muito além do aumento do espectro de proteínas presentes. Isto sugere o acúmulo de DNA sem significado

biológico no sentido de correspondência às proteínas (genes estruturais). Pode-se especular que a massa adicional de DNA tem função plástica, como constituinte dos cromossomos; funções regulatórias, controlando a expressão dos genes estruturais; não tem função definida, podendo aqui serem considerados genes que foram estruturais em antecessores evolutivos das espécies atualmente encontradas. Sabemos hoje que a expressão dos genes se dá através da síntese de um RNA, que difere do DNA apenas por conter uracila em vez de timida e ribose em vez de desoxirribose. Assim como a dupla fita de DNA é mantida por ligações denominadas Watson-Crick entre A e T e G e C especificamente, a fidelidade da expressão gênica é garantida pela cópia fiel do DNA em RNA. Assim, a seqüência de DNA ATGGCTAA é transcrita no Segmento RNA: UACCGAUU.

Como o genoma de organismos superiores comporta massa de DNA correspondente a algo como 10^5 genes, até o advento da Engenharia Genética parecia impossível conseguir-se isolar um gene a partir de tal diluição ($1:10^5$). É, porém, possível hoje construir genes sintéticos a partir do RNA mensageiro encontrado nas células, introduzir tais genes em bactérias mediante o uso de vetores adequados e determinar a seqüência destes genes após sua amplificação pela reprodução das bactérias. Dispomos assim de material perfeitamente caracterizado e em massa suficiente para buscar, no DNA celular, as seqüências gênicas correspondentes.

O isolamento dos primeiros genes permite verificar que ^(1, 2), anatômicamente, as seqüências estruturais não são contínuas isto é, co-lineares com o RNA delas aparentemente transcrito, mas interrompidas por seqüências de DNA que não tem correspondência com a proteína codificada. Estas últimas seqüências são chamadas *introns*, enquanto que as seqüências expressas são chamadas *exons*. Sabemos hoje que na verdade o primeiro produto de transcrição, denominado RNA hete-

* Professora Adjunta da Faculdade de Medicina da Fundação do ABC, Santo André, São Paulo.

** Professor Titular, Chefe do Laboratório de Oncologia Experimental da Faculdade de Medicina da Universidade de São Paulo, São Paulo.

rogêneo nuclear, é co-linear com o gene, ocorrendo processamento pós-transcricional deste produto, com renovação das seqüências correspondentes a introns e fusão de exons próximos. Tal construção dos genes poderia encontrar justificativa como meio de reduzir ou tamponar a carga mutacional. Desta forma, mutações ocorridas a nível de introns seriam silenciosas, não afetando a célula. Pode-se mostrar, por exemplo que há dois alelos para o gene da ovalbumina em galinha, que diferem entre si pela presença ou ausência de um sítio de restrição a nível de um intron⁽³⁾.

Outra possibilidade seria a construção de genes diferentes usando basicamente os mesmos exons. Assim, por exemplo, a diferença entre o gene para IgM de membrana e IgM secretada está em um único exon presente apenas no gene da IgM de membrana⁽⁴⁾.

Finalmente, podemos imaginar a utilidade dos introns como mecanismo para a fusão de exons distantes. Como a seqüência de nucleotídeos na fronteira exan-intron é constante, podemos juntar dois exons distantes durante a diferenciação celular. Por exemplo, a distância entre os exons que constituem o gene de uma imunoglobina é muito maior nos gametas do que no linfócito terminalmente diferenciado⁽⁵⁾.

Este novo conceito sobre a estrutura gênica surgiu após o advento de uma série de técnicas que facilitaram o isolamento de genes e sua caracterização. O conjunto dessas técnicas é conhecido como a *tecnologia de DNA recombinante* e, à ciência que as utiliza, têm-se o nome de *Engenharia Genética*.

I — A tecnologia do DNA recombinante: técnicas de clonagem molecular

De forma muito geral, a *clonagem molecular* consiste em ligar fragmentos de DNA "in vitro" e introduzi-los em células-vivas, onde tais moléculas possam ser replicadas, gerando indivíduos que contêm milhares de cópias idênticas ao fragmento original, sendo denominados *clones recombinantes* ou, simplesmente, clones.

A tecnologia, que vem sendo empregada nos mais diferentes sistemas, tanto animais como vegetais, se baseia em algumas descobertas independentes e que serão discutidas, em detalhe, mais adiante:

1. a descoberta, em bactérias, de endonucleases que cortam o DNA em locais específicos, chamadas enzimas de restrição^(6, 7, 8, 9).
2. o desenvolvimento de vetores moleculares para a clonagem, a partir de um conhecimento mais aprofundado dos plasmídeos bacterianos^(10, 11, 12, 13).
3. a descrição de um método para a introdução de DNA em bactérias^(14, 15).
4. as técnicas de seqüenciação de DNA, permitindo a caracterização de qualquer gene clonado^(16, 17, 18).

A clonagem molecular se reveste de uma importância ainda maior quando pensamos que o fragmento de DNA clonado é responsável pela síntese de uma proteína de interesse médico (hormônios, vacinas etc.).

Sob este aspecto, poder-se-ia empregar esse conjunto de técnicas para a produção em massa de diferentes produtos biológicos, como, de fato, vem ocorrendo há alguns anos.

Dada a pontencialidade dessa moderna tecnologia e o interesse geral que o assunto desperta, vamos descrever, com alguns detalhes, a metodologia normalmente empregada .

I.1 — As enzimas envolvidas na manipulação de genes "in vitro"

As *enzimas de restrição* são endonucleases capazes de reconhecer seqüências específicas na molécula de DNA e causar a quebra das duas fitas.

Na natureza, essas enzimas isoladas de bactérias, são responsáveis pela degradação de moléculas de DNA "estranhas" que, eventualmente, pudessem penetrar na célula.

Normalmente, as enzimas de restrição clivam o DNA entre os terminais 5'-fosfato e 3'-hidroxila e, por causa da simetria da seqüência de reconhecimento, geram moléculas contendo extremidades 5' fosfato proeminentes ou 3' fosfato complementares. Neste caso, diz-se que as enzimas geraram extremidades *coesivas* nas moléculas de DNA. Algumas enzimas pertencentes a essa classe, seguidas do microorganismo a partir do qual foram isoladas, são : Bam HI (*Bacillus amyloliquefaciens* H); Eco RI (*Escherichia coli* RY 13), Hind III (*Haemophilus influenzae* Rd), Pst I (*Providencia stuartii*), Sal I (*Streptomyces albus* G).

Outras enzimas, contudo, como por exemplo Hae III (*Haemophilus aegyptus*) ou Sma I (*Serratia marcescens* Sb) clivam a dupla fita no mesmo ponto, gerando extremidades perfeitamente pareadas ("Flush ends").

Deve-se ressaltar que já foram descritas mais de trezentas enzimas de restrição diferentes⁽¹⁹⁾, cada uma das quais reconhece uma seqüência específica de nucleotídeos no DNA. Portanto, o uso apropriado dessas enzimas permite o isolamento de fragmentos discretos de DNA, uniformes em tamanho e capacidade de codificação, uma vez que os cortes são realizados em regiões definidas.

Além disso, as referidas endonucleases são um importante instrumento para análise dos genes clonados, pois através delas pode-se estabelecer o mapa físico da seqüência gênica em questão.

A obtenção de fragmentos de DNA contendo seqüências complementares em suas extremidades, como

descrito acima, é a primeira etapa a ser vencida no processo de clonagem de um gene⁽²⁰⁾. Segue-se a *ligação* desses fragmentos, "in vitro", a vetores moleculares que se encarregarão de transportá-los para dentro de células e, assim, gerar clones. Tal reação de ligação é efetuada pela *DNA ligase*, enzima capaz de formar uma ligação covalente entre extremidades de DNA^(21, 22). As duas ligases mais utilizadas provêm de *E. coli* ou do fago T4. Ambas catalisam a mesma reação porém, a primeira requer NAD como cofator, enquanto a codificada pelo fago T4 é dependente de ATP.

Assim, podem-se ligar fragmentos de DNA digerido com Hind III, por exemplo, a vetores moleculares contendo as mesmas extremidades coesivas e as moléculas resultantes dessa ligação utilizadas para a clonagem.

A desvantagem de se ligar DNAs através de extremidades coesivas é pela alta frequência de ligação que ocorre com os vetores entre si. Como resultado, tem-se vários clones que não contém o DNA a ser estudado, mas, apenas o vetor recircularizado (vide adiante). Entretanto, este problema pode ser sanado pelo uso de *fosfatase alcalina* (de bactéria ou intestino de bezerro), enzima capaz de retirar os grupos fosfato das extremidades das moléculas de DNA^(23, 24), impedindo, assim, a ação da *DNA ligase*. Uma vez que o fragmento de DNA de interesse possui os fosfatos 5' terminais a ligase pode uní-lo ao vetor defosforilado, sendo esta ligação reparada "in vivo".

Talvez o principal obstáculo em se utilizar extremidades coesivas para ligação de fragmentos de DNA seja pelo fato que, muitas vezes, o DNA a ser estudado é susceptível de digestão pela enzima de restrição escolhida, causando, portanto, a interrupção da seqüência analisada. Em vista disso, foram desenvolvidos dois métodos através dos quais esse obstáculo pode ser vencido.

A *adição de extensões homopoliméricas*^(25, 26) consiste em tratar fragmentos de DNA dupla fita com a enzima transferase terminal, na presença de um dos quatro desoxiribonucleotídeos. Esta enzima adicionará às extremidades 3' das moléculas resíduos sucessivos do nucleotídeo escolhido, gerando, por exemplo, extensão de poli A (desoxiadenosina). Se o mesmo procedimento for efetuado com vetor linearizado (vide adiante) mais o desoxiribonucleotídeo adequado, nesse caso dTTP, serão geradas extremidades complementares capazes de serem ligadas ao fragmento de DNA de interesse, por ação do *DNA ligase*, com já descrito.

Extensões de poli A ou poli T são mais facilmente adicionadas pela transferase terminal, podendo-se chegar a 100 resíduos adicionados em cada extremidade. No caso de escolha de dCTP ou dGTP, tais nucleotídeos serão adicionados em menor número (aproximadamente 20 resíduos). Entretanto, o uso desse último par de nucleotídeos resulta na regeneração do sítio espe-

cífico para a enzima de restrição Pst I, permitindo que o fragmento de DNA inserido possa ser retirado do vetor, o que não ocorre quando se adicionam extensões de poli A: poli T.

Um outro método consiste na utilização de *fragmentos de DNA sintetizados quimicamente*, cujo tamanho varia entre 8 e 15 pares de nucleotídeos, contendo a seqüência de reconhecimento de uma determinada enzima de restrição. Quando esses fragmentos sintéticos são ligados às extremidades de um DNA qualquer, passam a exibir a capacidade de serem clonados, após digestão, no sítio correspondente ao de um vetor escolhido. Em outras palavras, o método leva à adição de um sítio de enzima de restrição à extremidade de um DNA que, anteriormente, não poderia ser facilmente clonado^(27, 28, 29).

Dentre as enzimas envolvidas na clonagem dos genes, merece destaque uma enzima naturalmente envolvida na replicação de vírus de RNA, capaz de sintetizar DNA, usando como molde uma fita de RNA. Esta enzima, a *transcriptase reversa*^(30, 31) pode ser utilizada para a obtenção de moléculas de DNA dupla fita, a partir de um RNA mensageiro (mRNA) isolado de células ou tecidos diferentes. Este método é aplicável, sobretudo, no caso de tecidos engajados preferencialmente na síntese de determinadas proteínas, contendo, portanto, uma quantidade razoável de mRNA, como no caso de glândulas ou tumores definidos. Dessa forma, vários clones foram construídos até hoje e, a partir deles, os genes foram localizados e analisados em detalhe.

1.2 — Vetores de Clonagem Molecular

Uma vez que a essência da tecnologia de DNA recombinante reside na propagação de moléculas de DNA de diferentes organismos em células hospedeiras, como bactérias por exemplo, faz-se mister a descrição de *vetores* moleculares capazes de propagar aquelas moléculas "in vivo".

Os vetores comumente empregados são derivados de plasmídeos bacterianos, bacteriófagos ou vírus animais e vegetais e, em razão da aplicação das técnicas de DNA recombinante ao estudo dessas moléculas, a cada dia são descritas novas e mais completas características desses vetores. Assim, inúmeros vetores de clonagem molecular são conhecidos, mas nos deteremos na apresentação de suas características principais, ressaltando aqueles mais frequentemente utilizados.

1.2.1 — Plasmídeos bacterianos

Plasmídeos são moléculas de DNA circulares, dupla fita, capazes de auto-replicação e de se manter eventualmente em células bacterianas. Além dos genes necessários à replicação dessas moléculas para as células-filhas,

a grande maioria dos plasmídeos codificam para uma série de diferentes funções de importância médica ou econômica, tais como, antibióticos, toxinas, resistência a metais pesados etc.

A manipulação genética dessas moléculas de ocorrência natural levou ao desenvolvimento de alguns vetores para a clonagem molecular ^(10, 32, 33).

Os aspectos gerais a serem considerados em relação a esses vetores residem, principalmente, no seguinte: a) o tipo de controle de replicação ao qual está sujeito; b) a existência de sítios únicos para enzimas de restrição; c) a presença de marcadores seletivos.

Alguns plasmídeos estão sob o controle de replicação dito rigoroso ⁽³⁴⁾ ("stringent") que faz com que eles existam em número pequeno dentro da célula. Por exemplo o primeiro plasmídeo usado para clonar um fragmento de DNA eucarótico ^(20, 35), o pSC101, é um vetor que se apresenta em apenas uma ou duas cópias por célula. Este fato gera um sério problema em relação à quantidade de DNA que pode ser recuperada após a clonagem; em outras palavras, o rendimento por massa de células transformadas será baixo.

Contudo, há plasmídeos, como o Col E1 ⁽³⁶⁾ que existem em quinze ou mais cópias por cromossomo bacteriano. Além disso, esse número pode ser elevado até 1.500 ou mais cópias por células, se adicionarmos cloranfenicol a uma cultura bacteriana na fase logarítmica de crescimento ⁽³⁷⁾. Nessas condições, o cromossomo bacteriano tem sua replicação inibida, enquanto que os plasmídeos continuam a replicar-se, podendo chegar até a 50% do DNA total, resultando num excepcional rendimento. Outra vantagem do Col E1 é que ele não é um plasmídeo auto-transferível, tornando-o um vetor seguro para a maioria dos experimentos de clonagem

Por estas razões, a maioria dos vetores de transferência-molecular utilizados hoje em dia são plasmídeos multicópia derivados de Col E1 ^(10, 38, 39).

Para ser usado como um vetor de clonagem, o plasmídeo deve conter sítios únicos para enzimas de restrição, nos quais poderão ser inseridos os fragmentos de DNA a serem clonados. Obviamente tal inserção não pode prejudicar as funções normais de replicação do vetor, mas, para facilitar a identificação dos novos plasmídeos, deve acarretar na inativação de um marcador seletivo. Por exemplo, o pSC 101 apresenta um único sítio para a enzima EcoRI; a clonagem neste sítio não interfere nem com a sua replicação nem com a expressão de seu único marcador, o gene para resistência à tetraciclina. Entretanto o pBR 322, possui ⁽³⁸⁾ dois genes para resistência aos antibióticos ampicilina e tetraciclina e sítios únicos para enzimas de restrição em ambas as seqüências. Assim, a clonagem no sítio de Hind III do pBR 322 resulta na inativação do marcador de tetra-

ciclina, mantendo intacto o de ampicilina. Esta propriedade permite que as bactérias transformadas com tais plasmídeos possam ser identificadas em meios de cultura contendo antibióticos e, portanto, facilmente diferenciadas daquelas que não contêm plasmídeos recombinantes.

Dentro desse espírito, um sem número de manipulações genéticas foi efetuado com diferentes plasmídeos, a fim de introduzir marcadores adequados e sítios para enzimas de restrição, mantendo contudo um tamanho final discreto, adequado à sua introdução, manutenção e segregação em diferentes células hospedeiras.

1.2.2 — Vetores derivados do bacteriófago λ

O vírus λ tem um genoma composto de aproximadamente 50 genes, envolto por uma carapaça proteica, capaz de penetrar em bactérias e se reproduzir.

Desses 50 genes, contudo, apenas a metade são essenciais ao crescimento e formação de placas de lise, restando uma região de aproximadamente 20 quilobases que não codifica para funções essenciais ^(13, 40). É neste fato, sobretudo, que se baseia a construção de vetores para clonagem molecular, oriundos da deleção dessas regiões pouco importantes deixando espaço para inserir seqüências exógenas.

A replicação do DNA de bacteriófago λ ocorre bidirecionalmente ⁽⁴¹⁾, através de um processo que depende de alguns produtos gênicos do fago, utilizando as funções de replicação da bactéria ⁽⁴²⁾. Ao final de um ciclo complexo de replicação, as moléculas são "empacotadas" na carapaça proteica, sendo que apenas aquelas que contêm 78 a 105% de tamanho do genoma selvagem originam partículas virais viáveis ⁽⁴³⁾. Devido à esta particularidade, foram desenvolvidos vetores capazes de abrigar tanto inserções pequenas quanto grandes. Assim, um fragmento pequeno de DNA, de no máximo algumas quilobases, pode ser clonado num vetor derivado de fago λ do tipo vetor de INSERÇÃO, que contém sítios únicos para enzimas de restrição, em regiões não essenciais de seu genoma. Será, entretanto, necessário utilizar um vetor de SUBSTITUIÇÃO, quando se quiser clonar fragmentos de DNA de tamanho superior a 10 quilobases. Nesse caso, parte do genoma não essencial do vetor é *substituído* pelo fragmento a ser clonado, podendo abrigar seqüências de até 20 quilobases de comprimento ^(13, 44).

Vantagens adicionais sobre o processo de transfecção de fagos recombinantes em células bacterianas podem ser obtidos utilizando-se a técnica de empacotamento dos DNAs virais "in vitro" ^(45, 46). Este procedimento emprega extratos celulares e DNA de fago linearizado, permitindo altas eficiências de transfecção.

Existem vários métodos para reconhecer os fagos recombinantes, entre os quais, a formação de placas de lise claras em vez de turvas⁽⁴⁷⁾, placas de lise incolores em lugar de azuis ou mesmo através de hidridização com sondas radioativas (vide adiante). Outros métodos se baseiam na inativação de certas funções gênicas⁽⁴⁸⁾ ou, ainda, na separação dos fagos normais daqueles recombinantes em gradientes de densidade, por causa das diferenças de tamanho existentes entre eles⁽⁴⁸⁾.

Os vetores derivados de fago λ tem sido amplamente utilizados na construção de *bancos de genes ou bibliotecas genômicas*. Nestes, fragmentos de DNA, obtidos por quebra mecânica ou digestão com enzimas de restrição, a partir do genoma total, são clonados em vetores de substituição que conterão inserções de, aproximadamente, 20 quilobases. A amplificação de tais bancos gênicos, seguida de análise por hibridização "in situ", permite localizar genes que, muitas vezes, estão representados apenas um em 100.000⁽⁴⁹⁾.

Hoje em dia várias bibliotecas genômicas estão disponíveis^(13, 29) e muitos genes foram descritos a partir de sua análise.

1.3 — Métodos de caracterização dos clones recombinantes

Existem, basicamente, dois métodos para caracterização dos vetores recombinantes que contêm a seqüência de DNA desejada: o primeiro, e mais comum deles, consiste na caracterização física desses DNAs, envolvendo mapas de restrição, testes de hibridização e determinação da seqüência de nucleotídeos. No segundo método, o gene é selecionado porque é expresso na célula hospedeira.

Tendo-se isolado uma quantidade considerável de DNA plasmidial pode-se proceder à sua digestão com um sem número de enzimas de restrição, a fim de se obter o seu *mapa de restrição*. Se se está estudando um gene conhecido, a simples comparação entre os padrões de digestão permitirá identificar o clone. No caso de seqüências não descritas, a construção do mapa de restrição permite detectar os sítios mais adequados na geração de fragmentos para seqüenciação, como veremos mais adiante.

A caracterização de DNAs por *hibridização* se baseia no princípio da complementariedade de bases dos ácidos nucleicos. Assim, clones podem ser identificados através testes de DNA utilizando *sondas* específicas. Estas consistem de fragmentos de DNA conhecidos, que são tornados radioativos pela ação da DNA polimerase I, na presença dos desoxinucleotídeos marcados com radioisótopos (por exemplo, p^{32} ,^(50, 51)). Em condições especiais nas quais as fitas de DNA são mantidas separadas umas das outras, ou seja, desnaturadas,

poderá ocorrer hibridização surgindo moléculas híbridas radioativas. A visualização desses eventos é feita por *autoradiografia*, na qual chapas de filme de raio X são impressionadas pela radiação emitida pelas moléculas híbridas⁽⁴⁹⁾.

Baseados nesses princípios, vários métodos de hibridização foram descritos e são amplamente utilizados em Engenharia Genética. Todos eles requerem a fixação do DNA a ser estudado sobre filtros de nitrocelulose, o que é facilmente conseguido após desnaturação e incubação a altas temperaturas. De acordo com a forma através da qual os DNAs são transferidos para os filtros, podemos citar três técnicas principais: a) a *hibridização em colônias*⁽⁵²⁾, na qual as colônias bacterianas selecionadas são crescidas sobre os filtros de nitrocelulose e, posteriormente, o seu DNA é fixado e hibridizado; b) a *hibridização de Southern*⁽⁵³⁾ consiste na transferência de DNAs, que foram previamente separados por eletroforese em gel de agarose ou acrilamida, diretamente para os filtros que serão, então, tratados conforme já descrito; e c) a *hibridização por pontos* ("dot"), na qual os DNAs purificados são aplicados diretamente sobre os filtros de nitrocelulose⁽⁵⁴⁾. Em todos os casos, a hibridização é efetuada em soluções desnaturantes, sob temperaturas controladas, na presença de uma sonda radioativa específica. Os filtros de nitrocelulose secos são, então, expostos a filmes de raio X que, após revelação, mostrarão o aparecimento de bandas ou pontos negros, indicando hibridização positiva.

Por analogia à hibridização de Southern também está descrita aquela de *Northern*^(55, 56), na qual moléculas de RNA são fixadas nos filtros, diferentemente do descrito acima. Contudo, a tecnologia é muito semelhante e tem sido amplamente empregada na caracterização de clones recombinantes.

É, contudo, a *seqüenciação de DNA*, a mais impressionante e elucidativa técnica descrita nos últimos anos. Através de sua aplicação, novos clones foram identificados, genes foram descritos em toda a sua extensão, e, sobretudo, começam-se a conhecer as seqüências envolvidas na regulação e expressão de diferentes genes. Por exemplo, a seqüência AAUAAA é sempre encontrada próxima à extremidade 3' dos mRNAs poliadenilados e tem sido considerada um sinal para poliadenilação e/ou término de transcrição⁽⁵⁷⁾. Mais recentemente, foram identificadas seqüências similares a TATAAT aproximadamente 30 nucleotídeos antes do sítio de iniciação de transcrição da maioria dos DNAs eucarióticos⁽⁵⁸⁾ e, portanto, parece ser um sinal para o início desse evento celular.

Quase concomitantemente, dois métodos de seqüenciação foram descritos e vem sendo largamente aplicados. O método de Maxam e Gilbert⁽¹⁶⁾, baseia-se

em submeter o DNA a reagentes químicos que causam a quebra das bases nitrogenadas de forma específica. Numa primeira etapa o DNA deve ser tornado radioativo pela marcação de suas extremidades 5' (ou 3') pelo emprego da polinucleotídeo quinase de T4⁽¹⁷⁾ na presença de nucleotídeos marcados no fósforo γ . Segue-se uma digestão das moléculas marcadas com uma segunda enzima de restrição, permitindo a purificação de fragmentos que serão submetidos às reações químicas específicas. Assim, o material é dividido em 4 ou 5 alíquotas e, na presença de reagentes como hidrazina, dimetilsulfato ou ácido fórmico, são quebradas as bases T ou C, G e A, respectivamente. Estas reações são controladas para que ocorra apenas uma quebra por molécula, o que causará a formação de uma série de fragmentos de tamanho gradativamente maiores, correspondendo à ocorrência de cada base naquele DNA. As alíquotas são, então, aplicadas em géis de poliacrilamida muito finos e longos que permitirão a separação dos inúmeros fragmentos pelo seu tamanho e, uma vez que os mesmos são radioativos, o gel é coberto com uma chapa de raio X. O autoradiograma revelará um sem número de bandas que poderão ser identificadas de acordo com a reação empregada e lidas ordenadamente da posição mais inferior do gel até o seu topo, obtendo-se seqüência de nucleotídeos daquele DNA.

O segundo método baseia-se num princípio diferente, mas resulta igualmente na determinação da estrutura primária do DNA. Sanger e seus colaboradores utilizaram fragmentos de DNA simples fita, que por ação da DNA polimerase podem ser facilmente copiados, a partir de um "primer" (ou iniciador), semelhante ao que ocorre na replicação do DNA. Contudo, ao invés dos quatro desoxinucleotídeos normais, emprega-se um deles na forma do didesoxinucleotídeo correspondente, o que causará a interrupção da cadeia, gerando fragmentos de diferentes tamanhos⁽¹⁸⁾. Além disso, para permitir a sua visualização, um dos desoxinucleotídeos usado na reação é radioativo, por exemplo α ATP³². A dificuldade maior deste método reside na obtenção dos fragmentos de DNA simples fita. Contudo, Groneborn e Messing^(59, 60) desenvolveram vários vetores derivados do bacteriófago M13 — mp7, mp8, mp9, mp10, mp11 — que são empregados para clonar o fragmento de DNA que se deseja seqüenciar. Uma vez identificados os clones, os vírus são crescidos numa pequena quantidade de meio de cultura. Como esses fagos apresentam, além da forma replicativa dupla fita, uma forma simples fita que é liberada pela célula hospedeira, pode-se, diretamente, recuperar esses vírus simples fita que contém uma inserção a ser seqüenciada. Utilizando-se, então, uma pequena seqüência de DNA simples fita ("primer") que pareia imediatamente adiante do início da inserção⁽⁶¹⁾, pode-se, pela ação da DNA polimerase (fragmento Klenow) sintetizar os fragmen-

tos complementares em quatro reações definidas, conforme descrito acima.

Ambos os métodos permitem a "leitura" de 150 a 200 nucleotídeos por gel de seqüência, o que conforme as condições e o tamanho do clone, levará, em pouco tempo, à determinação da seqüência de nucleotídeos do gene.

Por ser uma ciência relativamente nova e em franca expansão, é óbvio que a cada dia sejam descritos novos métodos, que permitem analisar as estruturas gênicas mais facilmente. Só neste aspecto existem inúmeras publicações que facilitam ao especialista a obtenção de novos resultados.

II — Expressão de genes clonados em células procarióticas e eucarióticas

Os mecanismos de expressão gênica nos organismos procarióticos diferem sobremaneira daqueles que ocorrem nos eucariotos. Contudo, vários exemplos há que demonstram ser possível expressar genes de outros organismos em células bacterianas. Para tanto, é preciso que as seqüências estranhas estejam sob o controle dos mecanismos de transcrição e tradução da célula hospedeira.

A compreensão destes requisitos levou à construção de vetores específicos chamados de *expressão* em que a seqüência a ser transcrita é inserida a jusante de um promotor bacteriano. Um promotor é uma seqüência de DNA que dirige a ligação da RNA polimerase ao DNA para que seja feito o mRNA. Se o promotor for "forte", mais moléculas de mRNA serão sintetizadas. Há também outras regiões importantes para a expressão, como a "Pribnow box"⁽⁶²⁾ localizada aproximadamente 10 nucleotídeos — antes do início de transcrição, assim como uma região a — 35 pares de bases⁽⁶³⁾.

Os promotores mais utilizados para a expressão de seqüências exógenas em *E.coli* são aqueles que, além de fortes, são reguláveis. Entre eles, tem sido muito utilizados os promotores dos operons *trp* e *lac*, envolvidos no metabolismo do triptofano e da lactose, respectivamente. Ambos podem ser regulados e induzidos na presença de determinados compostos, a fim de se conseguir o máximo de expressão⁽⁶⁴⁾. Além disso, recentemente foi descrito um vetor de expressão que contém uma combinação dos promotores *trp* e *lac* — promotor *tac* — e que se demonstrou ser bastante forte e regulável⁽⁶⁵⁾.

Para conseguir níveis elevados de expressão gênica em bactérias é preciso que também esteja presente o local de ligação aos ribossomos, para assegurar que o mRNA seja eficientemente traduzido. Esse sítio conhecido como seqüência de Shine Dalgarno⁽⁶⁶⁾.

inclui um codon de iniciação AUG e uma sequência de 3 a 9 nucleotídeos localizada 3 a 11 nucleotídeos antes desse codon. Uma das primeiras demonstrações de que esse sistema podia funcionar envolveu o pequeno hormônio humano, somatostatina⁽⁶⁷⁾. O método foi empregado para expressar diversos genes tendo-se percebido que, apesar dos controles gênicos serem adequados, não se detectavam níveis razoáveis da proteína exógena. Esse fato levou à observação de que a proteína estranha era rapidamente degradada na célula hospedeira, o que resultou na construção de outros vetores de expressão. Nestes, uma sequência gênica bacteriana é mantida de forma a produzir-se uma proteína híbrida contendo na sua extremidade amino terminal uma sequência proteica da bactéria e, na porção carboxiterminal, a proteína eucariótica. O exemplo clássico dessa abordagem é a expressão em bactérias do hormônio de crescimento humano⁽⁶⁸⁾. Além de dar maior estabilidade das proteínas sintetizadas dentro da bactéria, a expressão através de proteínas híbridas permite em alguns casos a exportação dessas proteínas para o exterior das células, com vantagens evidentes em relação à sua purificação. Um bom exemplo foi descrito por Villa-Komaroff e colaboradores⁽⁶⁹⁾ em relação à insulina de rato. A clonagem de sequência gênica do hormônio a jusante ao gene da penicilinase existente no vetor bacteriano resultou na formação da proteína híbrida penicilinase-proinsulina que pôde ser detectada e isolada do espaço periplasmático bacteriano.

Há que considerar, contudo, que a expressão dependerá de que as sequências estejam adequadamente posicionadas, na fase correta da leitura. Para tanto, alguns vetores tem sido desenvolvidos^(70, 71). Uma outra forma de superar essa dificuldade consiste no uso de exonucleases que, em determinadas condições, poderão retirar alguns nucleotídeos da sequência a ser clonada e expressa.

Não bastassem os inúmeros obstáculos expostos acima, há que salientar que os genes eucarióticos são caracterizados pela presença de introns, inexistentes nos organismos procarióticos, de sorte que a clonagem de sequências genômicas em bactérias não resultará em sua expressão. Por esta razão, grandes esforços tem sido dispendidos para o aprimoramento de vetores para clonagem em células eucarióticas, as quais são capazes de processar os produtos de transcrição até o mRNA definitivo.

O hospedeiro eucariótico mais simples e bem conhecido sob o ponto de vista genético e bioquímico é a levedura *Saccharomyces cerevisiae*. Vários trabalhos tem demonstrado a transformação de leveduras por diferentes genes, tanto pela técnica de protoplastos, quanto pelo uso de vetores derivados do plasmídeo 2 micron encontrado em muitas leveduras⁽⁷²⁾. Mais recentemente,

genes de *Drosophila melanogaster* foram eficientemente clonados em levedura⁽⁷³⁾, além de se ter obtido expressão de um gene de interferon humano⁽⁷⁴⁾. Entretanto, esse gene não contém introns. Quando se experimentou a expressão do gene de β globina de coelho, que apresenta introns, a levedura foi incapaz de transcrevê-lo corretamente⁽⁷⁵⁾.

Assim, partiu-se para o aprimoramento de vetores que pudessem transformar diretamente células superiores. Por analogia aos vetores empregados em bactéria, e como até o momento não se encontraram plasmídeos em células superiores, ampliaram-se os estudos em relação aos vírus animais. O pequeno vírus de macaco SV 40 tem sido muito estudado e empregado em Engenharia Genética. Este vírus se divide em células de macaco, causando lise e morte celular. Contudo, em outras células, ditas não permissivas, uma pequena porcentagem das células é transformada e, assim se mantém, por apresentar o DNA viral integrado de forma estável no cromossomo da célula hospedeira. Da mesma maneira que foram construídos os vetores derivados do fago lambda, como descrito anteriormente neste artigo, foram também descritos vários vírus defectivos derivados de SV 40 e que tem sido empregados na formação de células superiores⁽⁷⁶⁾. Alguns exemplos demonstram que tanto cDNA quanto o gene de β globina de coelho foi eficazmente expresso em células de macaco, comprovando que o sistema pode reconhecer e processar os sinais eucarióticos⁽⁷⁷⁾.

É importante ressaltar que, nos sistemas eucarióticos, devem ocorrer também modificações pós-traducionais, para que as proteínas sejam funcionais. Obviamente essas reações não ocorrerão em bactérias e, novamente, o sistema eucariótico demonstrou ser eficiente no caso da expressão do antígeno de superfície do vírus de hepatite B humano⁽⁷⁸⁾. O antígeno não apenas é secretado pelas células em cultura, mas se apresenta corretamente glicosilado e com a conformação nativa esperada. Espera-se, portanto, que o sistema seja adequado à produção de vacinas.

Em vista dessa potencialidade, outros vetores de vírus animais estão sendo desenvolvidos⁽⁷⁶⁾.

Também se verificou que a microinjeção de SV 40 nos núcleos de oócitos de *Xenopus laevis* resultou no processamento correto do DNA viral⁽⁷⁹⁾. Além disso, outros sistemas eucarióticos são eficientemente transcritos e traduzidos com a maquinaria dos oócitos. Apesar de um estudo mais aprofundado ser necessário, esta também poderá ser uma forma de análise de expressão gênica num futuro próximo.

Ainda nos caberia comentar sobre experimentos de transformação direta de células de mamífero em cultura por DNAs virais. O gene da timidina quinase (tk) do

vírus de herpes simples ou de DNA de mamífero pode ser transferido para células deficientes nesse enzima (tk^-) e os transformantes são selecionados pois readquirem a capacidade de crescer em meio contendo timidina⁽⁸⁰⁾. Dessa forma, sequências genômicas de β globina de coelho clonadas em fago λ foram introduzidas em células de camundongo tk^- ⁽⁸¹⁾. Tais sequências, integradas no cromossomo das células transformadas, foram eficientemente transcritas, evidenciando-se a remoção dos introns.

IV — Considerações finais

Através desse extenso relato técnico, fica claro que um grande e rápido avanço foi dado em direção ao conhecimento da estrutura gênica em diferentes níveis e organismos. Outrossim, esse avanço permitiu a abertura de novas perspectivas em diferentes áreas, como por exemplo a produção, já em escala comercial, de in-

ulina humana ou interferon obtidos à partir de Engenharia Genética.

Contudo, é ilusório prever que este avanço será bem sucedido em todos os campos, a menos que um grande esforço de investigação seja feito. Até hoje, nenhum exemplo de expressão gênica pôde ser diretamente generalizado, sendo os níveis muito diferentes de um gene para o outro.

No Brasil, alguns laboratórios, em nível acadêmico, detêm a tecnologia de DNA recombinante e vem se dedicando tanto a estudos básicos quanto aos aspectos aplicados dessas técnicas.

Uma análise sobre os genes eucarióticos clonados até hoje deveria ser objeto de outro artigo, dada a quantidade alcançada. Contudo, uma lista completa, com bibliografias, foi recentemente publicada por Davies⁽⁸²⁾.

SUMMARY: The essence of Genetic Engineering lies in the splicing of DNA fragments or genes from different organisms to that of plasmidial or viral vectors, in order to be introduced and propagated in bacteria or other suitable hosts. A number of techniques has been described for the purpose of structural gene analysis, for the study of the mechanisms involved in their expression or in its regulation, or still, to produce specific proteins of medical or commercial import. Here we present a brief overview of several methods collectively designated *recombinant DNA technology*.

KEY WORDS: Genetic Engineering, cloning, clones, DNA, RNA, genes, plasmids, viruses, restriction enzymes, hybridization, DNA sequencing, expression vectors.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Glover, D. M. & Hogness, D. S. — *Cell* 10:167, 1977.
2. Tilgham, S. M.; Tiemeier, D. C.; Seidman, J. G.; Peterlin, B. M.; Sullivan, M.; Maizel, J. V. & Leder, P. — *Proc. Natl. Acad. Sci. USA* 75:725, 1978.
3. Dugaiczky, A.; Woo, S. L. C.; Colbert, D. A.; Lai, E. C.; Mace Jr., M. L. & O'Malley, B. W. — *Proc. Natl. Acad. Sci. USA* 76:2253, 1979.
4. Ravetch, J. V.; Siebenlist, V.; Korsmeyer, S.; Waldmann, P. & Leder, P. — *Cell* 27:583, 1981.
5. Hozumi, N. & Tonegawa, S. — *Proc. Natl. Acad. Sci. USA* 73:5652, 1976.
6. Smith, H. O. & Wilcox, K. W. — *J. Mol. Biol.* 51:379, 1970.
7. Kelly Jr., T. J. & Smith, O. H. — *J. Mol. Biol.* 51:393, 1970.
8. Nathans, D. & Smith, O. H. — *Ann. Rev. Biochem.* 44:273, 1975.
9. Roberts, R. J. — *Gene* 4:183, 1978.
10. Bolivar, F. & Backman, K. — *Methods in Enzymology* vol. 68, pg. 245-267, R. Wu ed., Academic Press, 1979.
11. Kahn, M.; Kolter, R.; Thomas, C.; Figurski, D.; Meyer, R.; Remaut, E. & Helinski, D. R. — *Methods in Enzymology* vol. 68, p. 268-280, R. Wu ed., Academic Press, 1979.
12. Thompson, R. — *Genetic Engineering* vol. 3, pg. 1-52, R. Williamson ed., Academic Press, 1982.
13. Brammar, W. J. — *Genetic Engineering* vol. 3, pg. 53-83, R. Williamson ed., Academic Press, 1982.
14. Mandel, M. & Higa, A. — *J. Mol. Biol.* 53:159, 1970.
15. Lederberg, E. M. & Cohen, S. N. — *J. Bacteriol.* 119:1072, 1979.
16. Maxam, A. M. & Gilbert, W. — *Proc. Natl. Acad. Sci. USA* 74:560, 1977.
17. Maxam, A. M. & Gilbert, W. — *Methods in Enzymology* vol. 65, part 1, pg. 499, 1980.
18. Sanger, F.; Nicklen, S. & Coulson, A. R. — *Proc. Natl. Acad. Sci. USA* 74:5463, 1977.
19. Roberts, R. J. — *Methods in Enzymology* vol. 68, pg. 27-40, R. Wu ed., Academic Press, 1979.
20. Cohen, S. N.; Chang, A. C. Y.; Boyer, H. W. & Helling, R. B. — *Proc. Natl. Acad. Sci. USA* 70:3240, 1973.
21. Lehman, I. R. — *Science* 186:790, 1974.
22. Higgins, N. P. & Cozzarelli, N. R. — *Methods in Enzymology* vol. 68, pg. 50-71, R. Wu ed., Academic Press, 1979.
23. Seeburg, P. H.; Shine, J.; Marshall, J. A.; Baxter, J. D. & Goodman, H. M. — *Nature* 220:486, 1977.
24. Ullrich, A.; Shine, J.; Chirgwin, J.; Pictet, R.; Tischer, E.; Rutter, W. J. & Goodman, H. M. — *Science* 196:1313, 1977.
25. Lobban, P. & Kaiser, A. D. — *J. Mol. Biol.* 78:453, 1973.
26. Roychoudhury, R.; Jay, E. & Wu, R. — *Nucleic Acids Res.* 3:101, 1976.

27. Bahl, C. P.; Mariani, K. J.; Wu, R.; Stavinsky, J. & Narang, S. — *Gene* 1:81, 1977.
28. Scheller, R. H.; Dickerson, R. E.; Boyer, H. W.; Riggs, A. D. & Itakura, K. — *Science* 196:177, 1977.
29. Maniatis, T.; Hardison, R. C.; Lacy, E.; Lauer, J.; O'Connell, C.; Quon, D.; Sim, G. K. & Efstratiadis, A. — *Cell* 15:687, 1978.
30. Verma, I. M.; Temple, G. F.; Fan, H. & Baltimore, D. — *Nature New Biol.* 235:163, 1972.
31. Efstratiadis, A.; Kafatos, F. C.; Maxam, A. M. & Maniatis, T. — *Cell* 7:279, 1976.
32. Novick, R. P.; Clowes, R. C.; Cohen, S. N.; Curtis III, R.; Datta, N. & Falkow, S. — *Bacteriol. Rev.* 40:168, 1976.
33. Bernard, H. U. & Helinski, D. R. — *Genetic Engineering* vol. 2, pg. 133, J. K. Setlow e A. Hollaender eds., Plenum Press, 1980.
34. Timmis, K.; Cabello, F. & Cohen, S. N. — *Proc. Natl. Acad. Sci. USA* 71:4556, 1974.
35. Morrow, J. F.; Cohen, S. N.; Chang, A. C. Y.; Boyer, H. W.; Goodman, H. M. & Helling, R. B. — *Proc. Natl. Acad. Sci. USA* 71:1743, 1974.
36. Hersfield, V.; Boyer, H. W.; Chow, L. & Helinski, D. R. — *J. Bacteriol.* 126:447, 1976.
37. Clewell, D. B. — *J. Bacteriol.* 110:667, 1972.
38. Bolivar, F.; Rodriguez, R. L.; Greene, P. J.; Betlach, M. C.; Heynecker, H. L.; Boyer, H. W.; Crosa, J. H. & Falkow, S. — *Gene* 2:95, 1977.
39. Twigg, A. J. & Sheratt, D. — *Nature* 283:216, 1980.
40. *The Bacteriophage Lambda*. Hershey, A. D. ed., Cold Spring Harbor Lab., 1971.
41. Schnös, M. & Inman, R. B. — *J. Mol. Biol.* 51:61, 1970.
42. Denniston-Thompson, K.; Moore, D. D.; Kruger, K. E.; Furth, M. E. & Blattner, F. R. — *Science* 198:1051, 1977.
43. Weil, J.; Cunningham, R.; Martin III, R.; Mitchell, E. & Bolling, B. — *Virology* 50:373, 1973.
44. Williams, B. G. & Blattner, F. R. — *J. Virol.* 29:555, 1979.
45. Enquist, L. & Sternberg, N. — *Methods in Enzymology* vol. 68, pg. 281-298, R. Wu ed., Academic Press, 1979.
46. Hohn, B. — *Methods in Enzymology* vol. 68, pg. 299-309, R. Wu ed., Academic Press, 1979.
47. Murray, N. E.; Brammar, W. J. & Murray, K. — *Mol. Gen. Genet.* 150:53, 1977.
48. Philipsen, P.; Kramer, R. A. & Davis, R. W. — *J. Mol. Biol.* 123:371, 1978.
49. Maniatis, T.; Fritsch, E. F. & Sambrook, J. — *Molecular Cloning*, pg. 269-294, Cold Spring Harbor Lab., 1982.
50. Olson, K. & Harvey, C. — *Nucleic Acids Res.* 2:319, 1975.
51. Rigby, P. W. J.; Dieckman, M.; Rhodes, C. & Berg, P. — *J. Mol. Biol.* 113:237, 1977.
52. Grunstein, M. & Hogness, D. — *Proc. Natl. Acad. Sci. USA* 72:3961, 1975.
53. Southern, E. — *J. Mol. Biol.* 98:503, 1975.
54. Kafatos, F. C.; Jones, C. W. & Efstratiadis, A. — *Nucleic Acids Res.* 7:1541, 1979.
55. Thomas, P. S. — *Proc. Natl. Acad. Sci. USA* 77:5201, 1980.
56. McMaster, G. K. & Carmichael, G. G. — *Proc. Natl. Acad. Sci. USA* 74:4835, 1977.
57. Proudfoot, N. J. & Brownlee, G. G. — *Nature* 263:211, 1976.
58. Corden, J.; Wasyluk, B.; Buchwalder, A.; Sassone-Corsi, P.; Keding, C. & Chambon, P. — *Science* 209:1406, 1980.
59. Messing, J.; Groneborn, B.; Muller-Hill, B. & Hofschneider, P. H. — *Proc. Natl. Acad. Sci. USA* 74:3642, 1977.
60. Groneborn, B. & Messing, J. — *Nature* 272:375, 1978.
61. Messing, J.; Crea, R. & Seeburg, P. H. — *Nucleic Acids Res.* 9:309, 1981.
62. Pribnow, D. — *Proc. Natl. Acad. Sci. USA* 72:784, 1975.
63. Rosenberg, M. & Court, D. — *Ann. Rev. Genet.* 13:319, 1979.
64. Miller, J. H. & Reznikoff, W. S. — *The Operon*, Cold Spring Harbor Lab., 1972.
65. DeBoer, H. A.; Comstock, L. J.; Yansura, D. G. & Heynecker, H. L. — in *Promoter Structure and Function*, R. L. Rodriguez e M. J. Chamberlain eds., 1982.
66. Shine, J. & Delgarno, L. — *Nature* 285:34, 1975.
67. Itakura, K.; Hirose, T.; Crea, R.; Riggs, A. D.; Heynecker, H. L.; Bolivar, F. & Boyer, H. W. — *Science* 198:1056, 1977.
68. Goeddel, D. V.; Heynecker, H. L.; Hozumi, T.; Arentzen, R.; Itakura, K.; Yansura, D. G.; Ross, K. J.; Miozzari, G.; Crea, R. & Seeburg, P. H. — *Nature* 281:544, 1979.
69. Villa-Komaroff, L.; Efstratiadis, A.; Broome, S.; Lomedico, P.; Tizar, R.; Naber, S. P.; Chick, W. L. & Gilbert, W. — *Proc. Natl. Acad. Sci. USA* 75:3727, 1978.
70. Tacon, W.; Carey, N. & Emage, S. — *Mol. Gen. Genet.* 177:427, 1980.
71. Talmadge, K.; Stahl, S. & Gilbert, W. — *Proc. Natl. Acad. Sci. USA* 77:3369, 1980.
72. Beggs, J. D. — *Genetic Engineering* vol. 2, J. K. Setlow e A. Hollaender eds., Plenum Press, 1980.
73. Henikoff, S.; Tatchell, K.; Hall, B. D. & Nasmyth, K. — *Nature* 289:33, 1981.
74. Hitzeman, R. A.; Hagie, F. E.; Levine, H. L.; Goeddel, D. V.; Ammerer, G. & Hall, B. D. — *Nature* 293:717, 1981.
75. Beggs, J. D.; van den Berg, J.; van Ooyen, A. & Weissmann, C. — *Nature* 283:835, 1980.
76. Rigby, P. W. J. — *Genetic Engineering* vol. 3, pg. 83-141, R. Williamson ed., Academic Press, 1982.
77. Hamer, D. H. & Leder, P. — *Nature* 281:35, 1979.
78. Moriarty, A. M.; Hoyer, B. H.; Shih, J. W. K.; Gerin, J. L. & Hamer, D. H. — *Proc. Natl. Acad. Sci. USA* 78:2606, 1981.
79. DeRobertis, E. M. & Metz, J. E. — *Cell* 12:175, 1977.
80. Wigler, M.; Sweet, R.; Sim, G. K.; Wold, B.; Pellicer, A.; Lacy, E.; Maniatis, T.; Silverstein, S. & Axel, R. — *Cell* 16:777, 1979.
81. Wold, B.; Wigler, M.; Lacy, E.; Maniatis, T.; Silverstein, S. & Axel, R. — *Proc. Natl. Acad. Sci. USA* 76:5684, 1979.
82. Davies, K. E. — *Genetic Engineering* vol. 3, pg. 142-173, R. Williamson ed., Academic Press, 1982.